GENOMIC ANALYSIS OF *Paenibacillus durus* ATCC 35681 AND ITS TRANSCRIPTOMIC RESPONSE TO NITROGEN-FIXING AND NITROGEN-ENRICHED CONDITIONS

MARDANI ABDUL HALIM

UNIVERSITI SAINS MALAYSIA

2017

GENOMIC ANALYSIS OF *Paenibacillus durus* ATCC 35681 AND ITS TRANSCRIPTOMIC RESPONSE TO NITROGEN-FIXING AND NITROGEN-ENRICHED CONDITIONS

by

MARDANI ABDUL HALIM

Thesis submitted in fulfillment of the requirements
for the degree of
Doctor of Philosophy

April 2017

ACKNOWLEDGEMENT

I would like to thank my mother and my late father for rising me to become what I am today and I wish to extend my love and gratitude to both of them including my wife and my two beautiful kids that supports me since the beginning of this project.

I wish to express my deepest gratitude to my supervisor, Professor Nazalan Najimudin for his guidance and his inspiring character that always motivated me throughout the whole year. I would also like to thank Professor Razip Samian for sharing his ideas and advices.

I would like to express my sincere gratitude to my friends: AFM, MSCA, NAA, MMM, NFL, DVY, BOB, POW, KBJ, YCW, AYAR, SKS, RAAM, NNI, NAMR, PM, LSL, KWANI and KMUN for all the advices, assistants and also gossips. Finally, I would like to thank MyBrain15 for financial assistant.

TABLE OF CONTENTS

CHAPTER FOUR: TRANSCRIPTOMIC AND REAL-TIME EXPRESSION ANALYSIS OF *nifH* GENES

CHAPTER FIVE: DETERMINATION OF TRANSCRIPTION START SITE (TSS) OF *nifB* AND *nifH* GENES

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

%	Percentage
β	Beta
γ	Gamma
~	Approximately
$\times g$	Times gravity
bp	Base pair
°C	Degree Celsius
C	Carbon atom
h	Hour
g	Gram
g/L	Gram per liter
kDa	KiloDalton
kg	Kilogram
M	Molar
min	Minute
mg	Milligram
mg/mL	Milligram per milliliter
mL	Milliliter
mM	Millimolar
nm	Nanometer
rpm	Rotation per minute
µg	Microgram
µg/mL	Microgram per milliliter

μM	Micromolar
μL	Microliter
μm	Micrometer
v/v	Volume per volume
wt%	Weight percent
w/v	Weight per volume
w/w	Weight per weight

LIST OF ABBREVIATIONS

ARA	Acetylene reduction assay
ATCC	American Type Culture Collection
BLAST	Basic local alignment search tool
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dTTP	Deoxythymidine triphosphate
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidiun bromide
GC	Gas chromatography
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pH	Potential hydrogen

ANALISIS GENOMIK *Paenibacillus durus* ATCC 35681 DAN TINDAK BALAS TRANSKRIPTOMIKNYA TERHADAP KEADAAN PENGIKATAN NITROGEN DAN KEADAAN DIPERKAYA NITROGEN

ABSTRAK

Paenibacillus durus ATCC 35681 ialah diazotrof Gram-positif yang boleh mengikat nitrogen walaupun dalam kehadiran nitrogen tetap seperti nitrat dan ammonium. Ini dianggap unik kerana aktiviti nitrogen dalam kebanyakan diazotrof direncat oleh sebatian ini. Dalam kajian ini, genom ATCC 35681 telah berjaya dijujuk dan dianalisis. Ia mempunyai saiz lebih kurang 5.5 Mb dengan kandungan G + C 51%. Sebanyak 5233 gen telah dikenal pasti dan 4808 daripadanya ialah gen pengekodan protein. Sebanyak 6 gen *nifH* dan 4 *nifB* telah ditemui dalam genom ini. Satu kelompok utama gen *nif* yang terdiri daripada *nifB1H1DKENXHesA* telah diteliti dan ia mengkodkan untuk nitrogenase berasaskan molibdenum. Gen pengkodan untuk enzim nitrogenase alternatif juga telah dikenal pasti. ATCC 35681 tidak memiliki NifA/NifL dan berkemungkinan ia menggunakan sistem GlnR untuk mengawal atur proses pengikatan nitrogen. Analisis transcriptomik perbandingan menunjukkan bahawa kelompok gen *nif* utama ini telah dikawalatur menurun dalam keadaan diperkaya nitrogen relatif kepada keadaan kekurangan nitrogen. Gen pengkodan untuk pengangkut sulfat dan molibdenum serta besi telah diekspreskan dengan tinggi dalam keadaan kekurangan nitrogen. Analisis serentak tahap pengikatan nitrogen dan pengekspresan gen *nifH* menunjukkan bahawa dalam keadaan nitrogen diperkaya, aktiviti nitrogenase dikesan hanya pada hari 1

pertumbuhan. Walau bagaimanapun, semua gen *nifH* telah diekspres dalam keseluruhan hari 1 hingga hari 9. Dalam keadaan kekurangan nitrogen, aktiviti nitrogenase direkodkan dari hari 1 hingga hari 6 manakala gen *nifH* sambung pengekspresan sehingga hari 9. Walaupun ketiadaan aktiviti pengikatan nitrogen, kehadiran transkrip *nifH* menimbulkan kemungkinan sistem pengawalaturan pasca-transkripsi atau pasca-translasi. Penghujung 5' transkrip mRNA untuk *nifH1*, *nifB1* dan *nifB2* telah berjaya dipetakan tetapi tiada promoter yang jelas telah dikenal pasti. Walau bagaimanapun, jujukan promoter jenis SigN telah ditemui di kawasan hulu *nifB1*. Kawasan hulu *nifB2* pula menunjukkan promoter yang dikenali oleh SigE yang merupakan faktor sigma yang terlibat dalam sporulasi. Ini menjelaskan bagaimana pengikatan nitrogen boleh berlaku semasa peringkat lewat pertumbuhan dalam media kekurangan nitrogen. Ini juga menunjukkan bahawa pengikatan nitrogen dan sporulasi berhubung kait dalam ATCC 35681 mungkin untuk menjana sebatian berasaskan nitrogen yang cukup untuk memastikan penyempurnaan pembentukan spora.

GENOMIC ANALYSIS OF *Paenibacillus durus* ATCC 35681 AND ITS TRANSCRIPTOMIC RESPONSE TO NITROGEN-FIXING AND NITROGEN-ENRICHED CONDITIONS

ABSTRACT

Paenibacillus durus ATCC 35681 is a Gram-positive diazotroph which can fix nitrogen even in the presence of fixed nitrogen such as nitrate or ammonium. This is considered unique because nitrogen fixation activity in most diazotrophs are inhibited by these compounds. In this study, the genome of ATCC 35681 was successfully sequenced and analysed. It has a size of approximately 5.5 Mb with a G+C content of 51%. A total of 5233 genes were identified and 4808 of these were protein coding genes. A total of 6 *nifH* and 4 *nifB* genes were found in the genome. A major *nif* gene cluster consisting of *nifB1H1DKENXHesA* was observed and it codes for molybdenum-based nitrogenase. Genes coding for alternative nitrogenase enzymes were also identified. ATCC 35681 lacked NifA/NifL and the nitrogen fixation process is possibly regulated by GlnR, a global regulator of nitrogen metabolism system. Comparative transcriptomic analysis showed that the major *nif* gene cluster was down-regulated under nitrogen-enriched condition with relative to nitrogen-deficient condition. Genes coding for sulfate and molybdenum as well as iron transporters were highly expressed under nitrogen-deficient condition. Simultaneous analysis of levels of nitrogen fixation and expression of *nifH* genes showed that under nitrogen-enriched condition, nitrogenase activity was detected only on day 1 of growth. However, all the *nifH* genes were expressed throughout from day 1 until day 9. Under nitrogen-deficient condition, nitrogenase activities

were recorded from day 1 until day 6 while the *nifH* genes continued their expression until day 9. The presence of *nifH* transcripts despite the absence of nitrogen fixation activity raised the possibility of a post-transcriptional or post-translational regulation of the system. The 5'-ends of *nifH1*, *nifB1* and *nifB2* mRNA transcripts were successfully mapped but no distinct promoter was identified. However, a SigN-type promoter sequence was found in the upstream region of *nifB1*. The upstream region of *nifB2* showed a promoter recognisable by SigE, which is a sigma factor involved in sporulation. This explained how nitrogen fixation could occur during late stage of growth in nitrogen-deficient medium. This also indicated that nitrogen fixation and sporulation are linked in ATCC 35681, possibly to generate enough nitrogen-based compounds to ensure completion of spore formation.

CHAPTER ONE

INTRODUCTION

Nitrogen is a colorless gas that is present in the air as dinitrogen (N_2) and it is an essential component of all proteins and nucleic acids in every organism (Vitousek *et al*., 2002). Despite dominating the atmosphere (about 78% by volume), it is a non-reactive gas and all living things cannot utilize this form of nitrogen. This type of nitrogen must be changed to other chemical forms so that it can be biologically used. The N_2 gas has to be converted or 'fixed' to a more usable form through a process called nitrogen fixation (Boyd and Peters, 2013).

The ability to fix nitrogen is possessed by certain species of *Bacteria* and *Archaea* and these are called diazotrophs. Members of the group includes all subdivision of the Proteobacteria (Dixon and Kahn, 2004). Within *Archaea*, some methanogens such as *Methanobacteriales, Methanococcales*, and *Methanosarcinales* were observed to be able to fix nitrogen (Boyds and Peters, 2013).

All diazotrophs possess an enzyme complex that catalyzes the reduction of nitrogen (N_2) to ammonia (NH_3) is known as nitrogenase. This complex consists of two components named according to their metal composition. The first component is also known as dinitrogenase reductase and it consists of a homodimer of Fe proteins. The monomer is encoded by the *nifH* gene. This component functions as an ATP-dependent electron donor to another component consisting of a larger heterotetramer of molybdenum-iron proteins (MoFe protein). The heterotetramer contains 2

molecules each of 2 non-identical proteins encoded by the genes *nifD* and *nifK* (Robert *et al*., 1987).

Some of the nitrogen-fixing bacteria produce non-molybdenum nitrogenases in addition to the molybdenum enzyme. These alternative nitrogenase consist of either vanadium and iron or iron only to replace the molybdenum. An alternative nitrogenase will not be synthesized when molybdenum is present. They only function as backup to ensure that the nitrogen fixation process can still happen in the absence of molybdenum (Madigan and Martinko, 2006).

The most well studied nitrogen-fixing bacteria are from the Gram-negative group. For example, *Klebsiella pneumoniae* and *Azotobacter vinelandii* have been the most used model systems for analysis of the organization and regulation of nitrogen fixation genes. Research activities that exploit Gram-positive diazotrophs as models are relatively limited compared to their Gram-negative relatives (Klipp *et al*., 2004). Among the Gram-positive diazotrophs that were studied are the obligate anaerobe *Clostridium pasterianum* and the actinomycete *Frankia alni* and both demonstrated distinctly different *nif* genes organization and regulation patterns (Wang *et al*. 1988, Harriott *et al*., 1995). In terms of gene regulation, the Gram-negative *K. pneumoniae* possesses a well-known positive regulator protein identified as NifA that mediates the expression of the nitrogenase genes in response to various environmental signals such as oxygen and nitrogen. In contrast, the *nifA* gene is absent in the genome of the Gram-positive diazotroph *Frankia alni* (Harriott *et al*., 1995).

Paenibacillus durus, formerly known as *Bacillus azotofixans* and *Paenibacillus azotofixans*, is a Gram-positive sporulating bacterium isolated from

Brazilian soils and grass roots (Seldin *et al.*, 1982). Subsequently, they were also found within the rhizospheres of maize, sugarcane, wheat and sorghum (Seldin *et al.*, 1984, Rosado *et al.*, 1996, Rosado *et al.*, 1998).

The type strain *Paenibacillus durus* ATCC 35681 has been shown to have the ability to fix nitrogen even in the presence of fixed nitrogen such as nitrate (Seldin *et al.*, 1984, Mollica *et al.*, 1985). Southern blot hybridization analysis using heterologous *nifH* probes revealed the possible presence of multiple *nifH* homologs in the genome of ATCC35681 (Seldin *et al.*, 1989, Choo *et al.*, 2003). The presence of multiple *nifH* genes in a single organism not only raised questions on their evolution but also their contribution in the nitrogen fixation process under varying environmental conditions.

The thesis aims to explore the evolution of nitrogen fixation genes in ATCC 35681 and their regulations under different nitrogen conditions. In order the achieve this, the following approaches were performed:

1. Whole genome sequencing was completed to analyze the genomic composition of ATCC 35681.
2. The transcriptomic profiles of cultures grown under nitrogen-rich (+N) and nitrogen-free (-N) conditions were acquired in order to explain the ability of ATCC 35681 to fix nitrogen in the presence of nitrate.
3. The activities of nitrogenase enzymes were measured via acetylene reduction assay coupled with qPCR to determine the expression of every *nifH* gene under the above-stated conditions.
4. RNA ligase-mediated rapid amplification of cDNA ends (RLM RACE) was executed to observe the regulatory 5’ regions of the genes,

CHAPTER TWO

LITERATURE REVIEW

2.1 History of biological nitrogen fixation and diazotrophs

Nitrogen is the most abundant element on earth and it is essential for all organisms to sustain life (Ferguson, 1998, Smil, 2004). In the atmosphere, it exists in the form of dinitrogen (N_2) gas which cannot be metabolized by most organisms (MacKay and Fryzuk, 2004, Jia and Quadrelli, 2014). Most organisms obtain their nitrogen source in the fixed form of nitrate (NO_3) or ammonia (NH_3) which is a fixed form of nitrogen via biological nitrogen fixation (Canfield *et al.*, 2010, Thamdrup, 2012).

The first successful demonstration of the biological assimilation of atmospheric nitrogen (N_2) by plant legumes for their growth and development was performed by Hellriegel and Wilfarth in 1888 as reviewed by Quispel (Quispel, 1988). In the same year, Beijerinck isolated bacteria from root nodules that aided the nitrogen fixation process leading to the isolations of *Azotobacter chroococcum*, *Azotobacter agilis* and *Clostridium pasteurianum* (Hu and Ribbe, 2011).

Since then, biological nitrogen fixation has been shown to be a major route for atmospheric nitrogen to be converted into ammonia (Seefeldt *et al.*, 2009) and it thus plays a key role in the global nitrogen cycle (Hu and Ribbe, 2011). Unfortunately, only a small number of microorganisms are able to assimilate atmospheric nitrogen into usable form (Burns and Hardy, 1975). Microorganisms that have the ability to fix atmospheric nitrogen into usable forms are known as

diazotrophs (Puri *et al.*, 2015a, Puri *et al.*, 2015b). Among the identified diazotrophs are members of eubacteria, archaea, actinomycetes, cyanobacteria and all subdivisions of Proteobacteria (Dixon and Kahn, 2004). Four out of five orders of methanogenic archaea are diazotrophs (Boone *et al.*, 2001) however, no eukaryote was included in the diazotrophic list (Young, 2005).

Most diazotrophs are free living. Nevertheless some diazotrophs have the ability to form symbiotic relationship with other organisms. The likes of *Rhizobium*, *Bradyrhizobium* and *Azorhizobium* can form symbiotic relationship with leguminous plants besides being able to grow as free living (Clark *et al.*, 2009, van Zeijl *et al.*, 2015).

The introduction of the acetylene reduction method for the measurement of biological nitrogen fixation activity greatly facilitated studies on this process as well as isolations of novel nitrogen fixing microbes (Dilworth, 1966). In addition, the development of semi-solid nitrogen-free medium that mimicked the microaerophilic environment within the soil niches was developed to aid the isolation of anaerobic nitrogen fixing bacteria from plant roots. From there onwards, extensive research about nitrogen fixation has been conducted.

2.2 The nitrogen cycle

Nitrogen cycle is one of the major cycles of chemical elements in the environment. This cycle involves several steps.

First step is the nitrogen fixation which can occur through naturally energy intensive event such as lightning or biologically by diazotrophs. The overall biological nitrogen fixation process is energy intensive and is as follows:

$$N_2 + 8H^+ + 8e^- + 16ATP \longrightarrow 2NH_3 + H_2 + 16ADP + 16Pi$$

Next step is known as nitrification which is a dual process whereby ammonia is oxidized into nitrite (NO_2^-) followed by another oxidation step to nitrate (NO_3^-). This process can be achieved by nitrifying bacteria usually by species of the genera *Nitrosomonas, Nitrosococcus, Nitrobacter* and *Nitrococcus.*

(a) $NH_4^+ \longrightarrow NO_2^-$ (followed by) (b) $NO_2^- \longrightarrow NO_3^-$

Nitrates can be reduced into nitrogen gas by denitrifying bacteria in a process called denitrification, thus releasing nitrogen back into the atmosphere (Gallon and Chaplin, 1987). This process can be accomplished by several species of bacteria from various genera such as *Bacillus, Pseudomonas, Serratia* and *Achromobacter*.

$$NO_3^- \longrightarrow N_2$$

Ammonification is a process of decomposition of organic nitrogen compounds to produce ammonia. This process can be performed by various microbes like from the genus *Pseudomonas, Vibrio, Proteus, Serratia, Bacillus, and Clostridium* (Herbert, 1999). Figure 2.1 summarized the biological nitrogen fixation process.

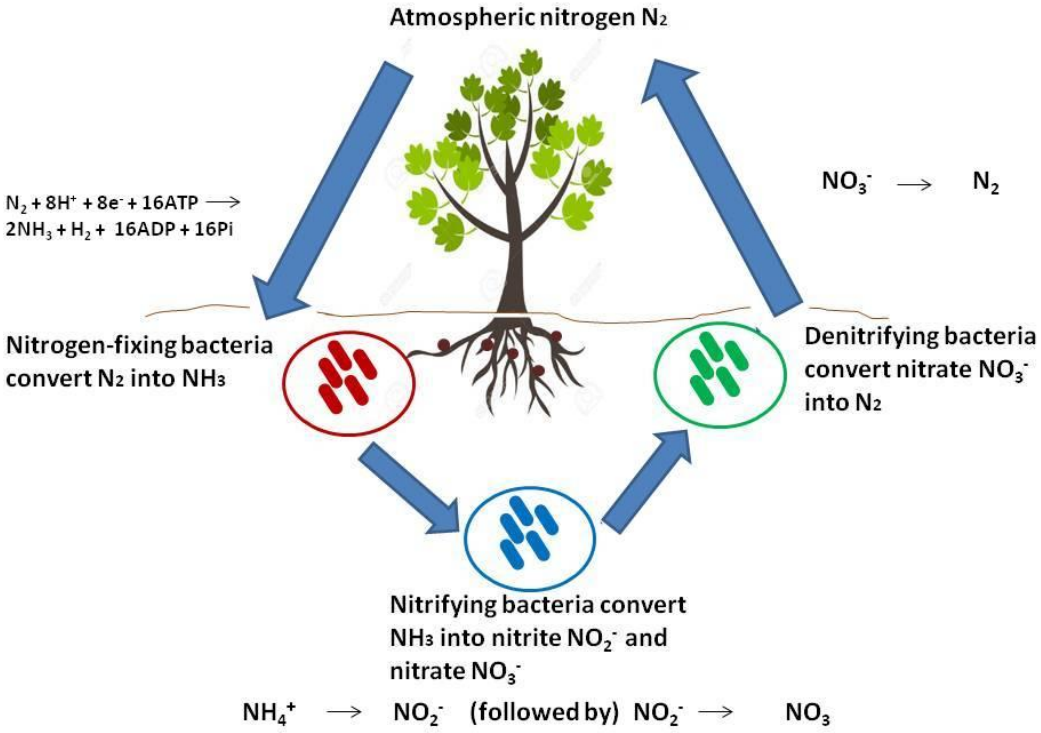

Figure 2.1 The nitrogen cycle. Nitrogen from the atmosphere is fixed by nitrogen-fixation bacteria into ammonia. Then, nitrifying bacteria will convert the ammonia into nitrite and nitrate. Finally, denitrifying bacteria will convert nitrate back into the atmosphere and the cycle continues.

2.3 Nitrogenase

Biological nitrogen fixation performed by diazotrophs is catalyzed by an enzyme complex known as nitrogenase and it is restricted to prokaryotes only (Teixeira *et al.*, 2008, Jiménez - Vicente *et al.*, 2015). Nitrogenase is a metalloenzyme which requires metal ion to catalyze the reduction of nitrogen (N_2) into ammonia (NH_3) (Seefeldt *et al.*, 2009). There are three types of nitrogenase system that have been identified, molybdenum-dependent nitrogenase (Mo), vanadium-dependent nitrogenase (V) and iron-dependent nitrogenase (Fe). These are encoded by *nif*, *vnf* and *anf* genes respectively (Dean and Jacobson, 1992, Boyd and Peters, 2013).

2.3.1 Molybdenum (Mo) nitrogenase

The Mo nitrogenase is the most characterized member within this enzyme family and has been successfully purified from *Azotobacter vinelandii*, *Clostridium pasteurianum,* and *Klebsiella pneumonia* (Lawson and Smith, 2002, Seefeldt *et al.*, 2009). This enzyme consists of two oxygen sensitive proteins designated as iron (Fe) protein and molybdenum-iron (MoFe) respectively (Seefeldt *et al.*, 2009) (Figure 2.2). Together, they catalyze the reduction of dinitrogen (N_2) with the following stoichiometry:

$$N^2 + 8e^- + 16MgATP + 8H^+ \rightarrow 2NH_3 + H^2 + 16MgADP + 16\ Pi$$

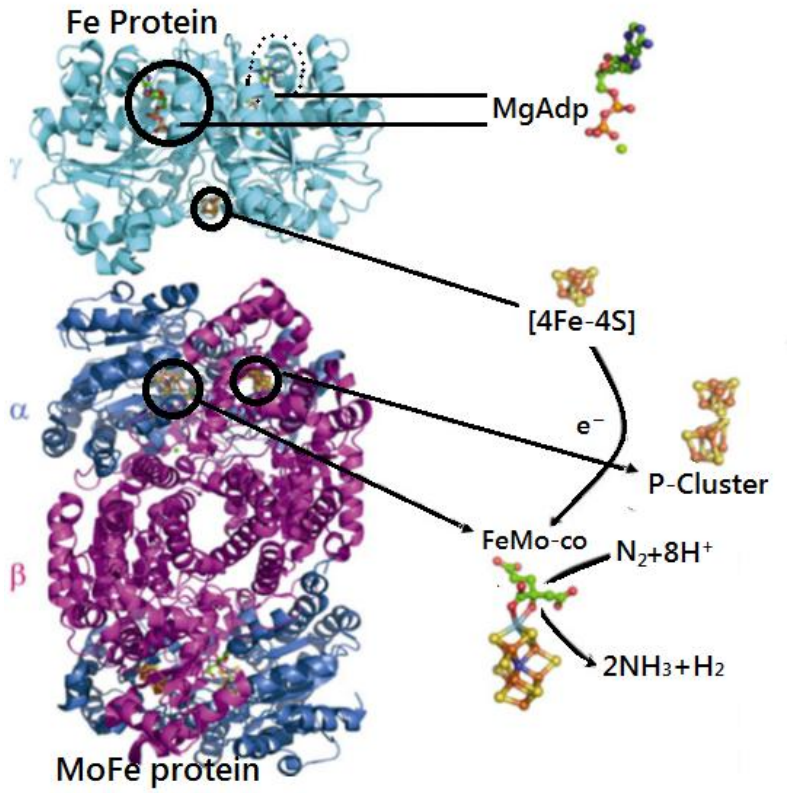

Figure 2.2 Crystal structure of the Fe and MoFe protein components of Mo-dependent nitrogenase. Left image represent Fe protein (pdb code: 1FP6) with the γ-subunits and MoFe protein (pdb code: 1M1N) with the α-subunits and β-subunits. (Adapted from Yang *et al*.,2011).

2.3.2 Fe protein

The Fe protein is an ATP-dependent reductase encoded by the gene *nifH*. It is a homodimer with a molecular weight of approximately 60 kDa (Strop *et al.*, 2001, Rubio and Ludden, 2008). Each of the homodimeric subunit contains a consensus motif (GXGXXG) which provides a binding site for MgATP. Besides providing electrons, it is also involved in the maturation of the MoFe protein and bioassembly of the active site metal cluster called MoFe cofactor (Rees *et al.*, 2005, Rubio and Ludden, 2008, Lee *et al.*, 2011). The two subunits of the Fe proteins are bridged by a [Fe4S4] cluster (Figure 2.2) through four Cys residues with two from each subunit (Rees *et al.*, 2005, Yang *et al.*, 2011).

During catalysis, the reduced Fe protein will bind to two MgATP molecules and subsequently associate with the FeMo protein (Hageman and Burris, 1978). Next, hydrolysis of MgATP to MgADP will occur and electron transfer takes place. Finally, the two proteins will dissociate and the used Fe protein will be reactivated by replacing the MgADP with MgATP (Yang *et al.*, 2011).

2.3.3 MoFe protein

The catalytic partner of the Fe protein in nitrogenase catalysis is the MoFe protein. It is a heterotetramer ($\alpha_2\beta_2$) with an approximate size of 250 kDa in which α and β subunits are encoded by *nifD* and *nifK*, respectively (Rubio and Ludden, 2008, Burgess and Lowe, 1996). Each αβ-dimeric unit contains one active metallocluster site formed by the FeMo cofactor ([7Fe-9S-Mo-X-(*R*)-homocitrate]) (Shah and Brill, 1977, Spatzal *et al.*, 2011) and one electron carrier cluster known as the P-cluster

([8Fe-7S]). Based on the X-ray crystal structure of the enzyme, the FeMo cofactor is embedded in the α-subunit while the P-cluster situated in between α and β-subunits (Figure 2.1) (Kim and Rees, 1992, Kim, 1993). Such an arrangement causes the FeMo protein to be treated as a dimer of dimers in which each of the αβ-dimer functions independently (Maritano *et al.*, 2001).

2.4 Alternative nitrogenase

Under the influence of certain growth condition, some of the nitrogen-fixing bacteria will produce non-molybdenum nitrogenases called alternative nitrogenases. This structure consists of either vanadium (V) and iron or iron (Fe) only to replace the molybdenum. For instance, this alternative nitrogenase system will not be utilized when the molybdenum is present. They only function as backups to ensure that the nitrogen fixation process can still happen when the molybdenum is absent (Clark *et al.*, 2009).

2.4.1 Vanadium (V) nitrogenase

The V nitrogenase in nitrogen fixation was initially identified as early as the Mo dependent enzyme but it was only verified during late 1980's. It was isolated from tungsten (W)-tolerant mutants, *A. chroococcum* (Robson et al., 1986, Eady et al., 1987, Eady et al., 1988) and also from *A. vinelandii* (Hales *et al.*, 1986a, Hales *et al.*, 1986b, Blanchard and Hales, 1996). Like Mo, the V nitrogenase also has two components, the Fe protein and vanadium-iron (VFe) protein (Eady, 1996, Hu *et al.*, 2012). Besides, the primary sequence for both Mo and V nitrogeanse are quite

similar. The *vnf* encoded Fe protein (*vnfH*) is also a homodimer with molecular weight ~60kD and it maintains the same conserved Cys ligands for the Fe_4S_4 cluster. On the other hand, the subunit of the VFe protein varies among species in which *A. chroococcum* has an $\alpha_2\beta_2\delta_2$ hexamer with molecular weight of ~240kDa (Eady, 1996) while *A. vinelandii* has an $\alpha_2\beta_2\delta_4$ octamer with a molecular weight of ~270kDa (Lee *et al.*, 2009).

2.4.2 Iron (Fe)-only nitrogenase

The iron-only nitrogenase is the least understood compared to the Mo and V nitrogenase. The enzyme was isolated from *Rhodospirillum rubrum*, *Rhodobacter capsulatus* and *A. vinelandii* (Eady, 1995, Eady, 1996). It is also a two component system consisting of Fe protein and iron-iron (FeFe) protein and the former is encoded by *anfH*. The α_2-homodimer subunits are bridged by a Fe_4S_4 cluster and the *anfH*-encoded Fe protein shows very close homology to the *nifH*- and *vnfH*- encoded Fe protein (Eady, 1996). Similar homology with Mo and V nitrogenases indicates that the Fe-only nitrogenase is likely utilizing the same mode of reaction as its counterparts (Hu and Ribbe, 2015).

2.5 Regulation of nitrogen fixation

Every diazotroph deals with physiological constraints during nitrogen fixation and they have evolved their own genetic and biochemical control systems in order to maximize the efficiency of the nitrogen fixation process. Some of the major and well addressed issues are the sensitivity of nitrogenase towards oxygen (O_2),

supply of substantial energy for the nitrogen fixing process, and also the metal requirement for the metallozymes (Merrick, 2005).

2.5.1 Transcriptional regulation of nitrogenase

To respond to the oxygen level and ammonium availability in the environment, the regulation of nitrogenase need to be done during the transcription level. It is essential for the bacteria to shut down the expression of metabolically expensive nitrogenase genes when the oxygen and fixed nitrogen levels are high (Merrick, 1992, Halbleib and Ludden, 2000).

Most studies on the transcriptional regulation of nitrogenase regulation is based on *K. pneumoniae*. The control of *nif* gene expression involved NifA and σ^{54} (*rpo*N gene product)-dependent transcriptional activator. The expression of *nif*A is under the control of the *ntrBC* gene products (*ntr*; nitrogen regulation system) which is a two component transcriptional activator system that is involved in the regulation of nitrogen in the cell (Merrick, 1992). NifA gene is cotranscribed with the redox and nitrogen-responsive regulatory flavoprotein (NifL) which is encoded by *nifL*. NifL also acts as negative regulator to NifA and thus serves another regulatory response to the fixed nitrogen and oxygen levels (Hill *et al.*, 1996).

The system employed by *K. pneumoniae* differs from one nitrogen fixer to another (Figure 2.3). For *A. vinelandii* and *R. rubrum*, the expression of *nifA* is not under the control of *ntrBC* gene products. In *Rhizobium meliloti*, a two component regulatory system (*fixL* and *fixJ*) is responsible for the expression of *nifA* in respond to the cell oxygen level (Merrick, 1992). *R. meliloti* also lacks NifL, but NifA is

present and is inhibited by oxygen. Such system replaces the *ntrBC* system employed by *K. pneumoniae* (Krey *et al.*, 1992). *R. capsulatus* has *nif*-related genes which are analogous to *ntrBC* but the expression of *rpoN*-like genes were found to be sensitive towards fixed nitrogen and oxygen levels (Klipp and Paschen, 1998).

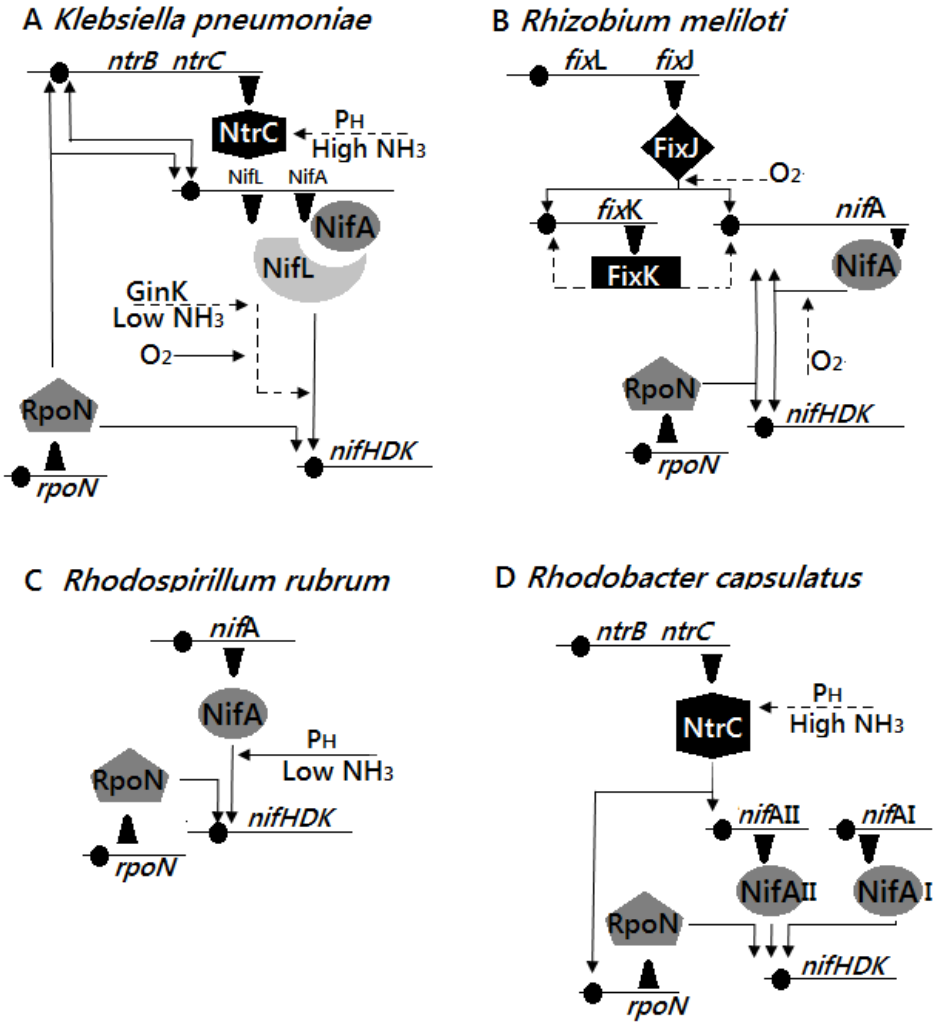

Figure 2.3 Transcriptional regulation of *nif* genes in *Klebsiella pneumoniae*, *Rhizobium meloliti*, *Rhodospirillum rubrum*, and *Rhodobacter capsulatus*. PII (the *glnB* gene product) and GlnK are global nitrogen regulation proteins of similar structure. RpoN is the sigma54 subunit of the RNA polymerase complex. Functionally similar regulatory proteins are symbolized identically in each panel. Solid and dashed lines indicate positive and negative regulation, respectively. Solid circles indicate promoter sites. Adapted from (Halbleib and Ludden, 2000).

2.6 Post translational regulation of nitrogenase

To control the regulation of nitrogen fixation when energy is limited or when nitrogen sources are available, an additional mechanism exists in which the nitrogenase complex is rapidly inactivated by ADP-ribosylation of the Fe protein. This system has been identified in *R. capsulatus*, *R. rubrum*, *Azospirillum brasilense* and *Chromatium vinosum* (Ludden and Roberts, 1989).

ADP-ribosylation of Fe protein in *R. rubrum* occurs at an arginine residue (Arg101) with the formation of an α-N-glycosidic bond between the arginine guanidino nitrogen atom and the terminal ribose of ADP-ribose (Ludden and Roberts, 1989). Dinitrogenase reductase ADP-ribosyltransferase (DRAT) and dinitrogenase reductase-activating glycohydrolase (DRAG) are encoded by *draT* and *draG*, respectively, and they are cotranscribed together with *draB*, a gene with an unknown function (Litwin *et al.*, 1992, Masepohl *et al.*, 1993). The gene *draB* may be involved in the regulation of nitrogenase activity by an unknown mechanism (Rubel *et al*., 2016).

DRAT, a monomer with a molecular weight of 30 kDa, has a high specificity towards oxidized MgADP-bound Fe protein, the activity of which cannot be measured (Ludden and Roberts, 1989, Halbleib *et al.*, 2000). DRAG is a monomer with a size of 32 kDa that catalyses the removal of ADP-ribose group and restores the activity of Fe protein with an intact arginine (Arg101) side chain. DRAG is also capable of cleaving the α-N-glycosidic bond from various ADP-ribosylarginine analogs (Ludden and Roberts, 1989). The exact mechanism of DRAT and DRAG is still unknown but, it is postulated that each will bind on the same surface of the Fe

protein (Grunwald *et al.*, 1995). Figure 2.4 summarized the reaction of DRAT and DRAG.

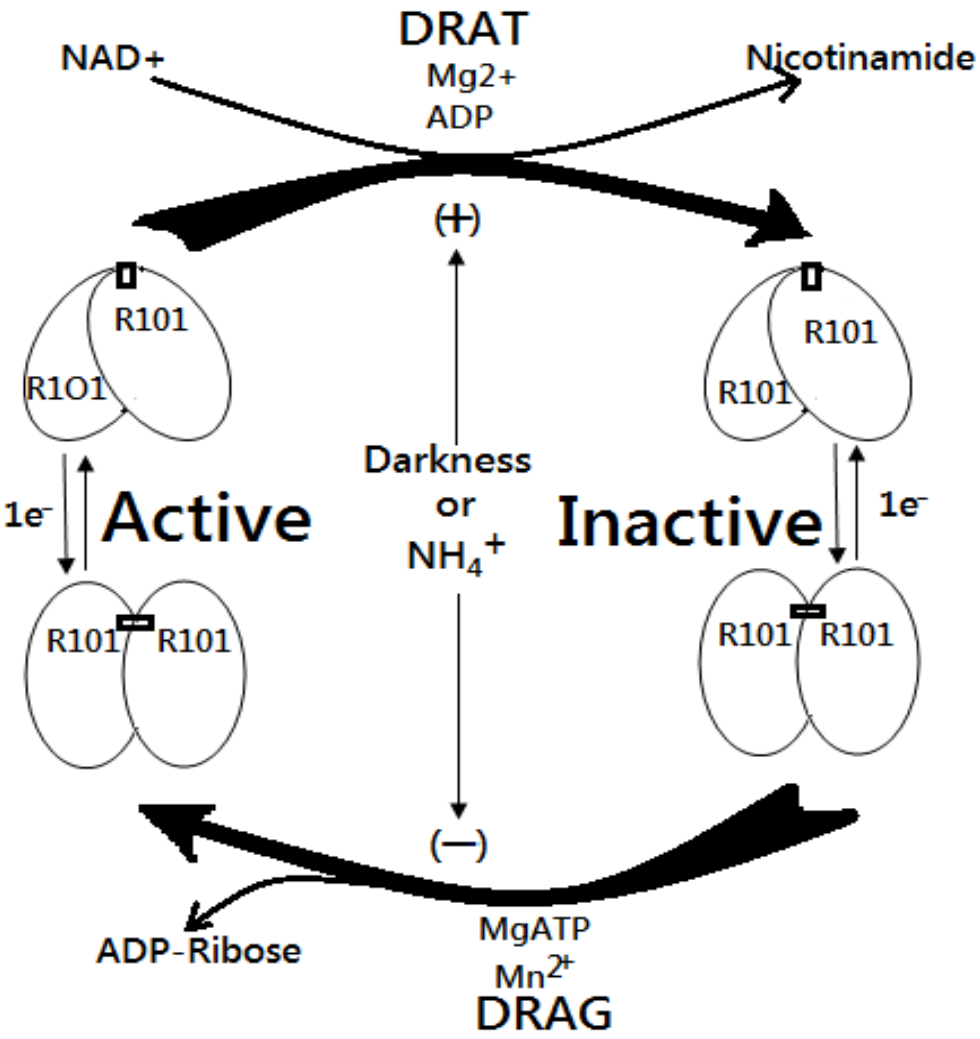

Figure 2.4 Model of *in vivo* nitrogenase Fe protein regulation in *R. rubrum* by reversible ADP-ribosylation. Small molecules required for dinitrogenase reductase ADP-ribosyltransferase (DRAT) and dinitrogenase reductase-activating glycohydrolase (DRAG) activities are shown besides the enzyme symbols. Adapted from (Halbleib and Ludden, 2000).

2.7 General nitrogen fixation control system

In general, all organisms regulate the nitrogen fixation process at the transcriptional level of the *nif* genes. Such regulation involves two stages: a general control system that regulate the metabolism of cellular nitrogen and a *nif*-specific regulation (Merrick, 2005, Vitousek *et al.*, 2013).

2.7.1 Nitrogen fixation in Proteobacteria

Most studies on the nitrogen regulation (Ntr) system were mostly done on the γ Proteobacteria but the system is also present in the α and β Proteobacteria. The Ntr regulation system involves four proteins: a PII signal transduction family, GlnB; uridylyltransferase, GlnD; sensor histidine kinase, NtrB and response regulator NtrC (Figure 2.5) (Masepohl, 2015, Merrick, 2005, Schumacher *et al.*, 2013). The Ntr system has been also found and characterized in *Azotobacter vinelandii*, *Klebsiella pneumoniae*, *Azospirillum brasilense*, *Rhodospirillum rubrum*, *Rhodobacter capsulatus*, *Gluconacetobacter diazotrophicus*, *Azoarcus* and *Herbaspirillum seropedicae* (Merrick, 2004, Peschek *et al.*, 2012, Nordlund, 2015).

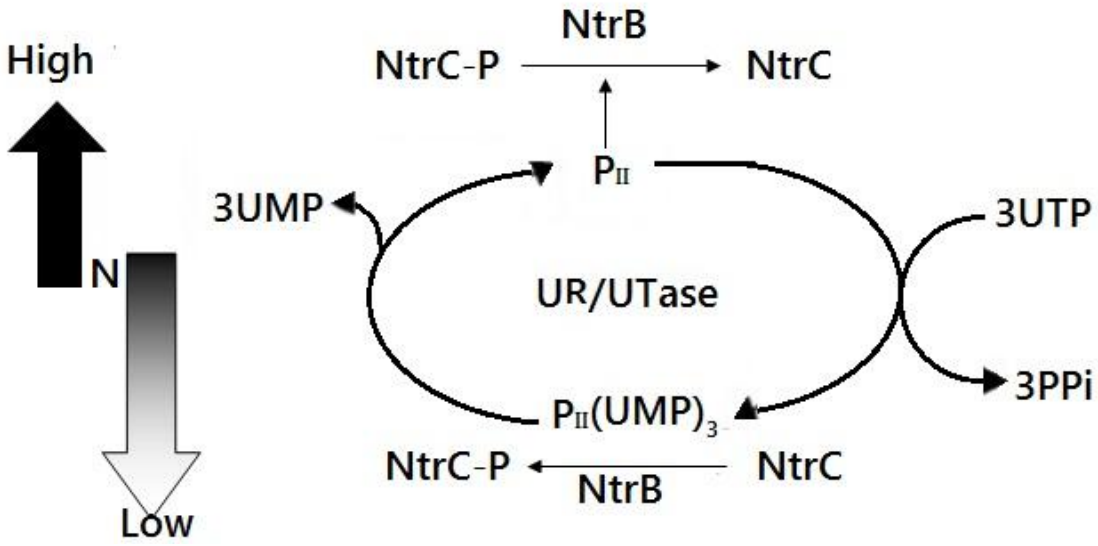

Figure 2.5 The nitrogen regulation (Ntr) system. The activity of the response regulator, NtrC, is regulated in response to the intracellular nitrogen status. UTase (*glnD* product) catalyses the uridylylation and deuridylylation of P_{11} (*glnB* product). The P_{11}, in turn, regulates the activity of the sensor histidine kinase NtrB, which catalyses the phosphorylation and dephosphorylation of NtrC. Adapted from (Merrick, 2005).

2.7.1(a) Nitrogen fixation in *Azotobacter vinelandii*

A. vinelandii is a free living nitrogen fixing bacterium and has been used as a model organism to study nitrogen fixation because (i) the nature of the genome of *A. vinelandii* that can be easily manipulated, (ii) it can fix nitrogen even in the presence of oxygen (~20%), (iii) nutritional flexibility, (iv) availability of its full genome sequence (Dingler *et al.*, 1988, Oelze, 2000, Setubal *et al.*, 2009). The main reason behind its ability to fix nitrogen even in the presence of oxygen is because of the availability of five terminal oxidases working together with its NADH oxidoreductases. The extra NADH oxidoreductases supply electrons to the terminal oxidases and this process consequently increases the cellular oxygen consumption (Setubal *et al.*, 2009).

A. vinelandii has three different nitrogenase; Mo-, V- and Fe-only that are located near the origin of replication of the chromosome (Setubal *et al.*, 2009). The genes of the Mo-dependent nitrogenase and together with their assembly and regulatory proteins system are located at two discrete regions. The main *nif*-region is around 25 kbp which include nine essential genes for nitrogen fixation (*nifH*, *nifD*, *nifK*, *nifE*, *nifN*, *nifU*, *nifS*, *nifV* and *nifM*) (Jacobson *et al.*, 1989a, Jacobson *et al.*, 1989b). All of the proposed functions of the essential and accessory *nif* genes are described in Table 2.1.

Table 2.1 ***A. vinelandii*** **genes involved in Mo-dependent nitrogen fixation** (Adapted from (O'Carroll and Dos Santos, 2011)).

Gene	Protein name	Function in nitrogen fixation
nifH	Fe protein, nitrogenase reductase	Provides electrons to nitrogenase and ATP-derivedenergy for catalysis
nifD	α-chain of nitrogenase or MoFe protein	Catalyzes reduction of nitrogen to ammonia
nifK	β-chain of nitrogenase or MoFe protein	Catalyzes reduction of nitrogen to ammonia
nifT		Unknown
nifY		Putative intermediate carrier of FeMo-cofactor
nifE		FeMo-cofactor assembly protein
nifN		FeMo-cofactor assembly protein
nifX		Putative intermediate carrier of FeMo-cofactor
FeSII	Shethna protein	Provides respiratory protection
$iscA^{Nif}$	IscANif	Involved in Fe–S cluster assembly protein
nifU	Fe–S cluster assembly scaffold protein	Assembles Fe–S clusters for nitrogen-fixing proteins
nifS	Cysteine desulfurase	Catalyzes transfer of sulfur from cysteine to NifU for Fe–S cluster formation
nifV	Homocitrate synthase	Required in biosynthesis of homocitrate moiety of FeMo-cofactor
cysE1	Serine *O*-acetyl transferase	Involved in biosynthesis of cysteine, which is used as a sulfur source of Fe–S cluster biosynthesis